On the Basic Representation Theorem for Convex Domination of Measures

J. Elton and T. P. Hill*

School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332-0160

metadata, citation and similar papers at core.ac.uk

A direct, constructive proof is given for the basic representation theorem for convex domination of measures. The proof is given in the finitistic case (purely atomic measures with a finite number of atoms), and a simple argument is then given to extend this result to the general case, including both probability measures and finite Borel measures on infinite-dimensional spaces. The infinite-dimensional case follows quickly from the finite-dimensional case with the use of the approximation property.

Key Words: fusion of a (probability) measure; dilation; convex domination; majorization; approximation property; Banach space; locally convex; topological vector space

1. INTRODUCTION

One of the basic theorems of convex domination is the result of Hardy et al. (1929, 1959) [5, 6], which says that if $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ are real numbers, then $\sum_{j=1}^n c(x_j) \geq \sum_{j=1}^n c(y_j)$ for all convex functions $c: \mathbb{R} \to \mathbb{R}$ if and only if there exists a doubly stochastic $n \times n$ matrix $M = (m_{ij})$ with $y_j = \sum_{k=1}^n m_{kj} x_k$ for all j (that is, $\underline{y} = M\underline{x}$, where $\underline{y} = (y_1, \ldots, y_n)^t$ and $\underline{x} = (x_1, \ldots, x_n)$). This basic result has been extended to probability measures on finite-dimensional spaces by Blackwell (1953) [1] and by Stein and Sherman (cf. [6]), to probability measures on various infinite-dimensional spaces by Cartier et al. (1964) [2] and Strassen (1965) [10], and to general

* Partially supported by U.S. National Science Foundation grants DMS-89-01267 and DMS-92-03524, a Fulbright research grant, and a Dutch National Science Foundation (NWO) grant, dossier B-61-281.

finite measures on $\mathbb{R}^1$ by Mirsky (1961) [8] and on infinite-dimensional spaces by Fischer and Holbrook (1980) [4], whose proof relied heavily on the Stein–Sherman theorem. The purpose of this note is twofold: to give an elementary geometric proof in the finitistic case (purely atomic with a finite number of atoms) in $\mathbb{R}^n$, in the spirit of the original result of Hardy et al. in $\mathbb{R}^1$, and to show how this elementary result can be used to easily give the general results with nonfinitistic measures and infinite dimensional spaces (separable Banach spaces or compact convex metrizable subsets of locally convex topological vector spaces).

Previous proofs of general cases have used various *ad hoc* arguments, and it seems not to have been noticed that all follow from the finitistic case. In particular, it should be of interest that the infinite-dimensional result follows quickly from the finite-dimensional case by an application of Grothendieck's approximation property. The language of *fusions* of measures, introduced in [3] for probability measures, will be used as the most natural setting for the proofs.

2. FINITE FUSIONS AND THE FISCHER–HOLBROOK THEOREM

Throughout this paper, *measure* will mean finite, nonnegative countably additive measure, and except for the last section, all measures will be (Borel) measures with finite support on finite-dimensional Euclidean space $\mathbb{R}^d$. For such a measure P, $\|P\|$ will denote its total mass $P(\mathbb{R}^d)$, $P(x)$ the P-measure of the singleton $\{x\} \in \mathbb{R}^d$, $supp\, P$ the support of P, and $b(P)$ the *barycenter* ($\|P\|^{-1}\int x\, dP(x) \in \mathbb{R}^d$) of P. The *Dirac delta measure* $\delta(x)$ for $x \in \mathbb{R}^d$ is $\delta(E) = 1$ if $x \in E$ and $= 0$ otherwise. For $z \in \mathbb{R}$, z^+ denotes the positive part $\max\{z, 0\}$ of z, and for $S \subseteq \mathbb{R}$, $|S|$ denotes the cardinality of S.

The next definition is a special case of Definition 3.5 in [3] of *fusion* for more general spaces and measures.

Definition 2.1. Suppose P and Q are measures in $\mathbb{R}^d$ with finite supports $supp\, P = X = \{x_1, \ldots, x_n\}$ and $supp\, Q = Y = \{y_1, \ldots, y_m\}$. Then Q is a *fusion* of P if there exists a nonnegative row-stochastic $n \times m$ matrix R satisfying

(i) $\vec{\mathbf{p}}\, R = \vec{\mathbf{q}}$ and

(ii) $\overrightarrow{\mathbf{px}}\, R = \overrightarrow{\mathbf{qy}}$,

where $\vec{\mathbf{p}} = (p_1, \ldots, p_n) = (P(x_1), \ldots, P(x_n)) \in \mathbb{R}^n$, $\overrightarrow{\mathbf{px}} = (p_1 x_1, \ldots, p_n x_n) \in \mathbb{R}^n$ (and similarly for $\vec{\mathbf{q}}$, $\overrightarrow{\mathbf{qy}}$).

Intuitively, the fusion Q is obtained from P via $R = (r_{ij})$ as follows. Start with P, which places mass p_i at $\{x_i\}$, $i = 1, \dots, n$. The first atom of Q, mass q_1 at y_1, is formed by removing fraction r_{i1} of the mass p_i at x_i for each $i = 1, \dots, n$, and fusing this total removed mass $q_1 = \sum_{i=1}^n r_{i1} p_i$ at the respective barycenter $y_1 = q_1^{-1} \sum_{i=1}^n r_{i1} p_i x_i \in \mathbb{R}^d$ (similarly for $q_2, y_2, \dots, q_m, y_m$).

(An alternative equivalent definition is that Q is a fusion of P if there is a nonnegative *column*-stochastic $n \times m$ matrix T with $\vec{\mathbf{y}} = \vec{\mathbf{x}}T$ and $T\overrightarrow{\mathbf{q}^t} = \overrightarrow{\mathbf{p}^t}$; the version in Definition 2.1 is chosen for symmetry and ease of intuitive description. For measures with finite mean (barycenter), Q is a fusion of P iff P is a *dilation* of Q; cf. [3].)

Let $\mathscr{C}$ denote the set of all nonnegative convex real-valued functions on $\mathbb{R}^d$.

DEFINITION 2.2. For two (finitely supported) measures P and Q on $\mathbb{R}^d$, *P convexly dominates Q* (written $P \succ Q$) if

$$\int c\, dP \geq \int c\, dQ \qquad \text{for all } c \in \mathscr{C}.$$

(An extension of this definition to more general measures and spaces and its equivalence to the definition in [3] for probability measures are contained in Section 5.)

The following theorem (conclusions (i) and (iii)) is the fusion version of the finite-dimensional Fischer–Holbrook (1980) result.

THEOREM 2.3. *Let P and Q be finite measures with finite supports in $\mathbb{R}^d$. Then the following are equivalent*:

(i) $\int c\, dP \geq \int c\, dQ$ *for all* $c \in \mathscr{C}$ (*i.e.*, $P \succ Q$).

(ii) $\int c\, dP \geq \int c\, dQ + (\|P\| - \|Q\|)^+ c(v)$ *for all* $c \in \mathscr{C}$.

(iii) *There exists a fusion $\tilde{P}$ of P that majorizes Q, i.e.*, $\tilde{P} \geq Q$.

(iv) $\hat{P} = Q + (\|P\| - \|Q\|)^+ \delta(v)$ *is a fusion of* P.

(*where* $v = (\|P\| - \|Q\|)^{-1}(\|P\| b(P) - \|Q\| b(Q)) \in \mathbb{R}^d$).

Observe that v is simply that point in $\mathbb{R}^d$ where the "excess" mass $(\|P\| - \|Q\|)$ must be placed to retain the barycenter of P.

The equivalent combinatorial or matrix-theoretic version of (i)–(iii) is as follows (the proof given below, however, will be in the above fusion setting).

THEOREM 2.3′. *Fix positive integers $n \geq m$, and let $\{x_1, \dots, x_n\}, \{y_1, \dots, y_m\}$ be finite subsets of $\mathbb{R}^d$. Then the following are equivalent*:

(i) $\sum_{i=1}^n c(x_i) \geq \sum_{j=1}^m c(y_j)$ *for all* $c \in \mathscr{C}$.

(ii) $\sum_{i=1}^n c(x_i) \geq \sum_{j=1}^m c(y_j) + (n - m) c((n - m)^{-1}(\sum_{i=1}^n x_i - \sum_{j=1}^m y_j))$ *for all* $c \in \mathscr{C}$.

(iii) *There exists a doubly stochastic* $n \times n$ *matrix* $M = (m_{ij})$ *with* $y_j = \sum_{i=1}^n x_i m_{ij}$ *for all* $j = 1, \ldots, n$, (*i.e.*, $\underline{y} = [M\underline{x}]_m$, *where* $\underline{y} = (y_1, \ldots, y_m)^t$, $\underline{x} = (x_1, \ldots, x_n)$, *and* $[\underline{v}]_m$ *is the first* m *components of the column vector* $\underline{v}$).

Remarks. The power of Theorem 2.3 and the key difference from the constant mass (probability measure) analog is the surprising "something-for-nothing" implication (i) ⇒ (ii), which is vacuous if P and Q have the same total masses. Given the set of inequalities (i), the stronger (recall $c \geq 0$) set of inequalities (ii) follows. This implication clearly may fail for *individual* c (i.e., $\int c\, dP \geq \int c\, dQ \nRightarrow \int c\, dP \geq \int c\, dQ + (\|P\| - \|Q\|)^+ c(v)$). It should also be noted that the class $\mathscr{C}$ can be replaced by the class of nonnegative convex *polyhedral* functions in the conclusion of the theorem, since P and Q have finite supports.

3. PROOF OF MAIN THEOREM

To begin with, two geometric lemmas will be established that will be key ingredients in the proof of the main theorem. The first will be used to construct a fusion that preserves the integral of a special convex function, and the second will be used to apply this technique to special points guaranteed to be in the domain of such fusions. Throughout this section, $co(X)$ is the closed convex hull of $X \subset \mathbb{R}^d$.

DEFINITION 3.1. For c: $\mathbb{R}^d \to \mathbb{R}$, and T a finite subset of $\mathbb{R}^d$, let $\breve{c} = \breve{c}_T$: $co(\mathrm{T}) \to \mathbb{R}$ be the function

$$\breve{c}(y) = \inf\{z \in \mathbb{R}: (y, z) \in co\{(x, c(x)): x \in T\}\} \qquad \text{for all } y \in co(T).$$

LEMMA 3.2. *Let* c: $\mathbb{R}^d \to \mathbb{R}$ *be convex, and let* T *be a finite subset of* $\mathbb{R}^d$. *Then*

(i) $\breve{c}$: $co(T) \to \mathbb{R}$ *is convex and piecewise affine*.

(ii) $\breve{c}(y) \geq c(y)$ *for all* $y \in co(T)$.

(iii) $\breve{c}(x) = c(x)$ *for all* $x \in T$.

(iv) *For each* $y \in co(T)$ *there exists a subset* S *of* T *and positive numbers* $\{\lambda_x\}_{x \in S}$ *so that* $\sum_{x \in S} \lambda_x = 1$, $\sum_{x \in S} x \lambda_x = y$, *and* $\sum_{x \in S} \breve{c}(x) \lambda_x = \breve{c}(y)$, *and* y *has a unique convex combination representation in* S (*that is, if* $\{\alpha_x\}_{x \in S}$ *are nonnegative with* $\sum_{x \in S} \alpha_x (1, x) = (1, y)$, *then* $\alpha_x = \lambda_x$ *for all* $x \in S$).

Proof. Observe that $\breve{c}$ is just the "lower" boundary of the convex polyhedron (in $\mathbb{R}^{d+1}$) that is the convex hull of the set $\{(x, c(x)): x \in T\}$.

Then (i)–(iii) follow easily since c is convex, and (iv) follows by projecting the lower face of K onto $\mathbb{R}^d$ and taking S to be the set of extreme points of the simplex of minimal dimension in $\mathbb{R}^d$ (formed from the projections of the extreme points of K) that contains y. ∎

LEMMA 3.3. *Let P and Q be finite measures in $\mathbb{R}^d$ with finite supports X and Y, respectively. If $P \succ Q$, then $co(X) \supset co(Y)$.*

Proof. It is enough to show that $co(X) \supset Y$. Let $y \in Y$, and suppose, by way of contradiction, that $y \notin co(X)$. By the basic separating hyperplane theorem, there is a hyperplane separating y and $co(X)$; that is, there is a linear functional $f\colon \mathbb{R}^d \to R$ and an $\alpha \in \mathbb{R}$ so that $f(y) > \alpha$ and $f(u) \le \alpha$ for all $u \in co(X)$. Letting $\phi\colon \mathbb{R}^d \to \mathbb{R}$ be given by $\phi(x) = \max\{f(x) - \alpha, 0\}$, observe that ϕ is convex (as the maximum of two affine functions) and nonnegative and satisfies $\phi(y) > 0$ and $\phi(u) = 0$ for all $u \in co(X) \supset \operatorname{supp} P$. But by definition of support, $Q(\{y\}) > 0$, so this implies that $\int \phi\, dQ > 0 = \int \phi\, dP$, contradicting the assumption $P \succ Q$. ∎

(Compare the analog of Lemma 3.3, Theorem 3.20 in [3], which asserts the corresponding inclusion of supports of the measures for more general spaces, but under the assumptions of fusion of two measures of the *same* mass. In addition, if P has a finite mean (barycenter), then convex domination is equivalent to fusion; this is the basic Theorem 4.1 of [3].)

It is easy to see that fusions *always* preserve both mass and barycenter. This and several other useful properties are recorded in the following proposition.

PROPOSITION 3.4. *Let P be a finite measure with support contained in a finite set $Z \subset \mathbb{R}^d$, and let $\mathscr{F}(P)$ denote the set of all fusions of P with support contained in Z. Then*

(i) $\|Q\| = \|P\|$ *for all* $Q \in \mathscr{F}(P)$.

(ii) $b(Q) = b(P)$ *for all* $Q \in \mathscr{F}(P)$.

(iii) $\mathscr{F}(P)$ *is compact and convex* (*when viewed as a subset of* $\mathbb{R}^{|Z|}$).

(iv) $\mathscr{F}(\mathscr{F}(P)) = \mathscr{F}(P)$, *that is, if* $Q \in \mathscr{F}(P)$ *and* $\hat{Q}$ *is a fusion of* Q (*with* $\operatorname{supp} \hat{Q} \subset Z$), *then* $\hat{Q} \in \mathscr{F}(P)$.

(v) *If* $Q \in \mathscr{F}(P)$, *then* $P \succ Q$.

Proof. Conclusions (i)–(iv) are straightforward from the definition of fusion, and (v) is an easy consequence of Jensen's inequality. (Alternatively, (iii) (convexity only), (iv), and (v) also follow from the more general infinite-dimensional versions in Theorem 3.11, Theorem 3.12, and Corollary 3.17, respectively, in [3].) ∎

For the remainder of this section, P and Q will be nonzero finite measures with finite supports X and Y, respectively, in $\mathbb{R}^d$. For $c\colon \mathbb{R}^d \to \mathbb{R}$, $P[c]$ denotes $\int c\, dP$. Let

$$m = \min\{P(x) \colon x \in X\}.$$

The main tools in this section can be expressed in terms of a certain easy special type of fusion, which will now be identified for ease of exposition.

DEFINITION 3.5. $\hat{P}$ is an *S-to-y P-fusion of mass transfer* $\hat{m} \le m$ if there exist nonnegative numbers $\{\lambda_x\}_{x \in S}$ so that $\sum_{x \in S} \lambda_x = 1$ and $\sum_{x \in S} x\lambda_x = y$ and so that

$$\begin{aligned}
\hat{P}(y) &= P(y) + \hat{m} \\
\hat{P}(x) &= P(x) - \hat{m}\lambda_x \qquad \text{for } x \in S \\
\hat{P}(x) &= P(x) \qquad \text{otherwise.}
\end{aligned}$$

(In other words, $\hat{P}$ takes mass only from S and places it all on a single point y, chosen so that the barycenter is preserved.)

The next three results form the basis for the proof of Theorem 2.3. The first lemma establishes the existence of a fusion of P preserving inequality of integral for a *given* $c \in \mathscr{C}$; the second is a trick using this single c to find a fusion of P that is uniformly "good" for *all* $c \in \mathscr{C}$; and the proposition builds on these to conclude the existence, for each $y \in Y$, of a fusion of P of strictly positive mass transfer that preserves convex domination of Q. Then the proposition is used via a minimality argument to establish the key implication in Theorem 2.3.

LEMMA 3.6. *Suppose $P \succ Q$ and $X \cap Y = \varnothing$. Given $y \in Y$ and $c \in \mathscr{C}$, there is a subset S of X such that y has a unique convex combination representation in S, and such that there is an S-to-y P-fusion P_S of mass transfer m such that $P_S[c] \ge Q[c]$.*

Proof. Fix $y \in Y$, $c \in \mathscr{C}$. Let $\check{c} = \check{c}_X$ be as in Definition 3.1. By Lemma 3.3, $y \in co(X)$, so by Lemma 3.2(iv) there is a subset S of X such that y has a unique convex combination representation $y = \sum_{x \in S} x\lambda_x$ and

$$\check{c}(y) = \sum_{x \in S} \check{c}(x)\lambda_x. \tag{1}$$

Let P_S be the S-to-y P-fusion of mass transfer m determined by $\{\lambda_x\}_{x \in S}$, that is,

$$\begin{aligned} P_S(Y) = m, \quad P_S(x) &= P(x) - m\lambda_x \qquad \text{for } x \in S, \\ P_S &= P \qquad \text{otherwise}. \end{aligned} \tag{2}$$

Since $supp\, P \cup supp\, Q \subset co(X)$,

$$P_S[\check{c}] = P[\check{c}] \geq Q[\check{c}], \tag{3}$$

where the equality follows from (1) and (2), and the inequality by the convexity of $\check{c}$ (Lemma 3.2(i)) and the hypothesis that $P \succ Q$. (Note $\check{c}$ is not actually defined off $co(X)$, but since $\check{c}$ is the maximum of a finite number of affine functions on $co(X)$, it has an immediate extension to a convex function on all $\mathbb{R}^d$.)

Next observe that

$$\begin{aligned} P_S[\check{c}] - P_S[c] &= (\check{c}(y) - c(y))P_S(y) \leq (\check{c}(y) - c(y))Q(y) \\ &\leq Q[\check{c}] - Q[c], \end{aligned} \tag{4}$$

where the equality follows by the definition (2) of P_S, and since $\check{c} = c$ on X (Lemma 3.2(iii)); the first inequality by (2) and the definition of m, since $P_S(y) = m \leq Q(y)$; and the last inequality since $\check{c} \geq c$ (Lemma 3.2(ii)). Together, (3) and (4) imply $P_S[c] \geq Q[c]$. ∎

LEMMA 3.7. *Suppose $P \succ Q$ and $X \cap Y = \varnothing$. Given $y \in Y$, there is an X-to-y P-fusion P_1 of mass transfer m satisfying*

$$Q[c] - P_1[c] \leq 2^{|X|}(P[c] - Q[c]) \qquad \textit{for all } c \in \mathscr{C}. \tag{5}$$

Proof. Fix $y \in Y$, and recall $y \in co(X)$ by Lemma 3.3. In fact, it will even be shown that for some $S \subset X$ there is an S-to-y P-fusion of mass transfer m satisfying (5) *and* such that y has a unique convex combination representation $y = \sum_{x \in S} x\lambda_x$ for some $S \subset X$. Suppose, by way of contradiction, that there is no such fusion. That is, for every subset S of X for which y has a *unique* representation $y = \sum_{x \in S} x\lambda_x$ there exists a $c_S \in \mathscr{C}$ so that if P_S is the unique S-to-y P-fusion of mass transfer m, then

$$Q[c_S] - P_S[c_S] > 2^{|X|}(P[c_S] - Q[c_S]). \tag{6}$$

Let $\mathscr{S} = \{S \subset X\colon y$ has a unique representation $y = \sum_{x \in S} x\lambda_x\}$, and let

$$\mathscr{S}_1 = \{S \in \mathscr{S}\colon P[c_S] > Q[c_S]\}$$

and

$$\mathscr{S}_2 = \{S \in \mathscr{S}\colon P[c_S] = Q[c_S]\}.$$

Note that $\mathscr{S}_1$ and $\mathscr{S}_2$ are disjoint, and since $P \succ Q$,

$$P[c_S] \geq Q[c_S] \qquad \text{for all } S \in \mathscr{S}, \tag{7}$$

so $\mathscr{S} = \mathscr{S}_1 \cup \mathscr{S}_2$.

Define $c \in \mathscr{C}$ by

$$c = \sum_{S \in \mathscr{S}_1} \frac{c_S}{P[c_S] - Q[c_S]} + \sum_{S \in \mathscr{S}_2} \frac{2^{|X|} c_S}{Q[c_S] - P_S[c_S]}.$$

(To see that $c \in \mathscr{C}$, note that c is the sum of positively weighted functions $c_S \in \mathscr{C}$, using the definition of $\mathscr{S}_1$ for the first sum, and (6) and (7) for the second.)

Since $c \in \mathscr{C}$, by Lemma 3.6 there is a subset S_0 of X such that y has a unique convex combination representation in $co(S)$ and so there is an S-to-y P-fusion of mass transfer m with

$$P_{S_0}[c] \geq Q[c]. \tag{8}$$

Observe that

$$Q[c] - P_{S_0}[c] = \sum_{S \in \mathscr{S}_1} \frac{Q[c_S] - P_{S_0}[c_S]}{P[c_S] - Q[c_S]} + \sum_{S \in \mathscr{S}_2} \frac{2^{|X|}(Q[c_S] - P_{S_0}[c_S])}{Q[c_S] - P_S[c_S]}. \tag{9}$$

Now,

$$P_{S_0}[c_S] \leq P[c_S] \tag{10}$$

by Proposition 3.4(v) (with $Z = X \cup Y$), since P_{S_0} is a fusion of P, and $P[c_S] = Q[c_S]$ for $S \in \mathscr{S}_2$, so $P_{S_0}[c_S] \leq Q[c_S]$ for $S \in \mathscr{S}_2$. This implies that the last summation in (9) is nonnegative. By (10) each term in the first summation in (9) is ≥ -1.

Case 1. $S_0 \in \mathscr{S}_1$. By (9) and (6),

$$Q[c] - P_{S_0}[c] \geq \sum_{S \in \mathscr{S}_1} \frac{Q[c_S] - P_{S_0}[c_S]}{P[c_S] - Q[c_S]} > -1(|\mathscr{S}_1| - 1) + 2^{|X|},$$

which is > 0, contradicting (8).

Case 2. $S_0 \in \mathscr{S}_2$. Similarly,

$$Q[c] - P_{S_0}[c] \geq -|\mathscr{S}_1| + 2^{|X|},$$

which is > 0, contradicting (8). ∎

PROPOSITION 3.8. *Suppose* $P \succ Q$ *and let* $X^* = \{x \in X\colon P(x) > Q(x)\}$ *and* $Y^* = \{y \in Y\colon Q(y) > P(y)\}$. *Given* $y \in Y^*$, *there exists an* X^**-to-*y P*-fusion* $\hat{P}$ *of strictly positive mass transfer such that* $\hat{P} \succ Q$.

Proof. First assume $X \cap Y = \varnothing$. Let P_1 be as in Lemma 3.7, so P_1 satisfies (5) and $P_1[c] = P[c] + mc(y) - m\Sigma_{x \in X} c(x)\lambda_x = P[c] - m\alpha_c$ for all $c \in \mathscr{C}$, where $\{\lambda_x\}_{x \in X}$ are nonnegative, $\Sigma_{x \in X}\lambda_x = 1$, $\Sigma_{x \in X} x\lambda_x = y$, and $\alpha_c = \Sigma_{x \in X}\lambda_x c(x) - c(y) \geq 0$, since c is convex. By (5),

$$m\alpha_c = P[c] - P_1[c] = P[c] - Q[c] + Q[c] - P_1[c]$$
$$\leq (2^{|X|} + 1)(P[c] - Q[c]) \qquad \text{for all } c \in \mathscr{C},$$

so

$$P[c] - Q[c] \geq (2^{|X|} + 1)^{-1} m\alpha_c > 0 \qquad \text{for all } c \in \mathscr{C}. \tag{11}$$

Letting $\hat{P}$ be the X-to-y P-fusion determined by the *same* $\{\lambda_x\}$, but mass transfer $\hat{m} = (2^{|X|} + 1)^{-1}m$, then

$$\hat{P}[c] - Q[c] = P[c] + \hat{m}c(y) - \hat{m}\sum_{x \in X} c(x)\lambda_x - Q[c]$$
$$= P[c] - Q[c] - \hat{m}\alpha_c \geq 0 \qquad \text{for all } c \in \mathscr{C},$$

where the inequality follows by (11). Now for the general case where $X \cap Y \neq \varnothing$, replace P by $P - (P \wedge Q)$ and Q by $Q - (P \wedge Q)$. ▮

Proof of Theorem 2.3. (i) ⇒ (iii). Suppose $P \succ Q$, and let $\mathscr{F}$ be the collection of all fusions $\tilde{P}$ of P satisfying

$$supp\,\tilde{P} \subset X \cup Y \quad (= supp\,P \cup sup\,Q) \tag{12}$$

and

$$\tilde{P} \succ Q. \tag{13}$$

Let

$$\gamma = \inf_{\tilde{P} \in \mathscr{F}} \left\{ \max_{y \in Y} \{Q(y) - \tilde{P}(y)\} \right\}.$$

Since X and Y are finite sets, and the set of fusions of P with support contained in $X \cup Y$ is closed (Proposition 3.4 with $Z = X \cup Y$), and since $\mathscr{F}$ is nonempty (since $P \in \mathscr{F}$), γ is attained. That is, there is a $\tilde{P} \in \mathscr{F}$ such that $\gamma = \max_{y \in Y}\{Q(y) - \tilde{P}(y)\}$. Without loss of generality, it may also be assumed that $|\{y \in Y\colon Q(y) - \tilde{P}(y) = \gamma\}|$ is minimal. It will now be shown that $\gamma \leq 0$, which establishes (iii).

Suppose, by way of contradiction, that $\gamma > 0$, and fix $\tilde{y} \in Y$ with $Q(\tilde{y}) - \tilde{P}(\tilde{y}) = \gamma$. Let $\tilde{X}^* = \{x \in \tilde{X}: \tilde{P}(x) > Q(x)\}$ and let $\tilde{Y}^* = \{y \in Y: Q(y) > \tilde{P}(y)\}$, where $\tilde{X} = supp\,\tilde{P}$. By (13) and Proposition 3.8 (applied to $\tilde{P}, \tilde{y}$ in place of P, y), there is an $\tilde{X}^*$-to-$\tilde{y}$ $\tilde{P}$-fusion $\hat{P}$ of strictly positive mass transfer $\hat{m}$ with $\hat{P} \succ Q$, such that $Q(\tilde{y}) - \hat{P}(\tilde{y}) = \gamma - \hat{m}$, and $Q(y) - \hat{P}(y) = Q(y) - \tilde{P}(y)$ for all other $y \in \tilde{Y}^*$ (since such y are not in $\tilde{X}^*$, and so their weights remain unchanged by an $\tilde{X}^*$-to-$\tilde{y}$ fusion). But this contradicts the minimality of $\tilde{P}$, so $\gamma \leq 0$.

(ii) ⇒ (i). Trivial, since $c \geq 0$ for all $c \in \mathscr{C}$.

(iv) ⇒ (ii). Since $\hat{P}$ is a fusion of P, $\hat{P} \succ P$ by Proposition 3.4(v) with $Z = X \cup Y \cup \{v\}$.

(iii) ⇒ (iv). $\hat{P}$ is the fusion of $\tilde{P}$ obtained by fusing all of the mass in $\tilde{P} - Q$. By Proposition 3.4(iv) (with $Z = X \cup Y \cup \{v\}$ again), $\hat{P}$ is a fusion of P, since it is a fusion of a fusion of P. ∎

4. EXTENSIONS TO GENERAL MEASURES AND INFINITE-DIMENSIONAL SPACES

The purpose of this section is to show how the basic finitistic (finite atoms, finite dimensions) result of Theorem 2.3 can be used to give simple proofs of analogous results in infinite-dimensional settings with general measures. Throughout this section, P and Q are finite Borel measures on V, where V is a separable Banach space or a compact convex subset of a locally convex topological vector space. Restriction to such spaces is only to ensure that *barycenters* exist; see [3] for details, as well as for the inclusion of *continuous* in the next definition.

DEFINITION 4.1. *P convexly dominates Q* (written $P \succ Q$) if $\int \phi\, dP \geq \int \phi\, dQ$ for all nonnegative continuous convex functions $\phi: V \to \mathbb{R}$ for which both integrals exist.

Remarks. Note that $P \succ Q \Rightarrow \int dP \geq \int dQ$, so $\|P\| \geq \|Q\|$. Also note that this definition agrees with Definition 3.15 in [3] (where *nonnegative* was not required) in case P and Q are probability measures, as is seen by the following argument: since $\int dP = \int dQ$, nonnegative convex domination implies $\int \phi\, dP \geq \int \phi\, dQ$ for all continuous convex functions that are bounded below. Then letting $\phi_t = \max\{\phi, -t\}$,

$$\int \phi\, dP = \lim_{t \to \infty} \int \phi_t\, dP \geq \lim_{t \to \infty} \int \phi_t\, dQ = \int \phi\, dQ.$$

The more general definition of *fusion* (Definition 2.1 above) for nonfinitistic probability measures and infinite-dimensional spaces given in [3] carries over easily to arbitrary positive finite measures. Intuitively, a fusion is simply the weak limit of measures formed from a base measure by repeatedly collapsing parts of the mass of measurable sets to their respective barycenters (see [3] for details).

The next theorem is the extension of Theorem 2.3 to general measures on infinite-dimensional spaces. For the identical-mass (probability measure) special case in infinite-dimensions, this gives a simple new proof of the main conclusions in Theorem 4.1 of [3] and of classical results in [2] and [9].

THEOREM 4.2. *Suppose P and Q are finite Borel measures on V, where V is a separable Banach space or compact metrizable convex subset of a locally convex topological vector space. If P has finite first moment (barycenter), then P convexly dominates Q if and only if there is a fusion $\hat{P}$ of P with $\hat{P} \geq Q$.*

The proof will be facilitated by several preliminary definitions and lemmas.

DEFINITION 4.3. Let V be a separable Banach space or a convex compact metrizable subset of a lctvs. In the case where V is a separable Banach space, assume that P has a finite first moment, that is, $\int \|x\| \, dP(x) < \infty$ (in the case where V is a convex compact metrizable subset of a lctvs, P will always *be said to have a finite first moment*). If A is a Borel set in V and $P(A) > 0$, then $b = b(A, P)$, the *P-barycenter* of A, is defined to be the unique element of the closed convex hull of A satisfying $f(b) = (\int_A f \, dP)/P(A)$ for all continuous linear functionals f on V (see [3], p. 422).

DEFINITION 4.4. A measure is *finitistic* iff it is purely atomic with finitely many atoms.

DEFINITION 4.5. Let V be a separable Banach space or a convex compact metrizable subset of a lctvs. If (t_{ij}) is an $n \times k$ row-stochastic matrix with nonnegative entries and A_i, $i = 1, \ldots, n$, is a Borel partition of V with $t_{ij} = 0$ if $b(A_i, P)$ does not exist, then the finitistic measure

$$\sum_{j=1}^{k} \sum_{i=1}^{n} t_{ij} P(A_i) \delta(a_j), \quad \text{where } a_j = \sum_{i=1}^{n} t_{ij} P(A_i) b(A_i, P) \Big/ \sum_{i=1}^{n} t_{ij} P(A_i)$$

is called a *finitistic matrix simple fusion* of P, written $\mathrm{fus}((A_i); (t_{ij}); P)$.

DEFINITION 4.6. The finitistic measure $\sum_{j=1}^{k} r_j \delta(z_j)$ is said to be an *ϵ-perturbation* of $\sum_{j=1}^{k} q_j \delta(y_j)$ if $\mathrm{dist}(y_j, z_j) < \epsilon$ and $q_j(1 - \epsilon) \leq r_j \leq q_j(1 + \epsilon)$ for $j = 1, \ldots, k$.

LEMMA 4.7. *Let V be a separable Banach space or a convex compact metrizable subset of a lctvs. Let P convexly dominate Q, where P is a finite positive Borel measure with a finite first moment, and Q is a finitistic positive measure $\sum_{j=1}^{k} q_j \delta(y_j)$ Then for all $\epsilon > 0$, there is a finitistic matrix simple fusion of P that is an ϵ-perturbation of a measure majorizing Q.*

Proof. *Part* 1. First assume V is a D-dimensional Euclidean space, and that the diameter of the support of P is finite.

Let $\epsilon > 0$. Cover *supp* P with finitely many simplices $\{S_j\}_{j=1}^{m}$ with vertices $\{v_i\}_{i=1}^{n}$ such that $\operatorname{diam}(S_j) < \epsilon$ and $P(\partial S_j) = 0\ \forall j$, and the S_j have nonoverlapping interiors. Let s_j be the P-barycenter of S_j. Write $s_j = \sum_i \alpha_{ji} v_i$, where $\alpha_{ji} = 0$ unless v_i is an extreme point of S_j, and $\sum_i \alpha_{ji} = 1$, $\alpha_{ji} \geq 0$. Let $m_j = P(S_j)$, let $p_i = \sum_j \alpha_{ji} m_j$, and let $P^* = \sum_i p_i \delta(v_i) = \sum_i \sum_j \alpha_{ji} m_j \delta(v_i)$, which is finitistic. Thus $P^*(\mathbb{R}^D) = \sum_i p_i = \sum_j m_j = P(\mathbb{R}^D)$.

Let $c\colon \mathbb{R}^D \to \mathbb{R}$ be a nonnegative convex function. For each j, let a_j be the affine function such that $a_j(v_i) = c(v_i)$ for each vertex v_i of S_j. Then $a_j(x) \geq c(x)\ \forall x \in S_j$, so

$$\int c\,dQ \leq \int c\,dP \leq \sum_j \int_{S_j} a_j\,dP = \sum_j m_j a_j(s_j) \qquad \text{(because } a_j \text{ is affine)}$$

$$= \sum_j m_j a_j\Big(\sum_i \alpha_{ji} v_i\Big) = \sum_i \sum_j m_j \alpha_{ji} c(v_i)$$

$$= \sum_i p_i c(v_i) = \int c\,dP^*.$$

Thus by Theorem 2.3, there is a fusion $\tilde{Q}$ of P^* that majorizes Q, so there exists a *row-stochastic* $n \times k$ matrix $\{t_{ij}\}$ such that

$$\tilde{Q} = \sum_{\ell=1}^{k} \left(\sum_{i=1}^{n} t_{i\ell} p_i \right) \delta(b_\ell),$$

where $b_\ell = \sum_i t_{i\ell} p_i v_i / \sum_i t_{i\ell} p_i$ and $\tilde{Q} \geq Q$. Let $u_{j\ell} = \sum_i \alpha_{ji} t_{i\ell}$, $j = 1, \ldots, m$, $\ell = 1, \ldots, k$. Note that $\{u_{jl}\}$ is also *row-stochastic*:

$$\sum_\ell u_{j\ell} = \sum_i \sum_\ell \alpha_{ji} t_{i\ell} = \sum_i \alpha_{ji} = 1.$$

Consider the fusion of P:

$$\tilde{\tilde{Q}} = \sum_{\ell=1}^{k} \left(\sum_{j=1}^{m} u_{j\ell} m_j \right) \delta(a_\ell),$$

where

$$a_{\ell} = \frac{\sum_j u_{j\ell} m_j s_j}{\sum_j u_{j\ell} m_j}.$$

(This is a fusion of P since the measure $\sum m_j \delta(s_j)$ is a fusion of P.)

Now since $\sum_j \alpha_{ji} m_j = p_i$,

$$\sum_j \alpha_{ji} m_j s_j = p_i \sum_j \left(\frac{\alpha_{ji} m_j}{p_i} \right) s_j = p_i v_i^*,$$

where $d(v_i, v_i^*) < \epsilon$, since v_i^* is a convex combination of $\{s_j\}$ with $d(s_j, v_i) < \epsilon$ (recall that $\alpha_{ji} = 0$ unless $d(s_j, v_i) < \epsilon$). Note also that $\sum_j u_{j\ell} m_j = \sum_j \sum_i \alpha_{ji} t_{i\ell} m_j = \sum_i t_{i\ell} p_i$. Note also that

$$\sum_j \sum_{\ell} u_{j\ell} m_j \delta(s_j) = \sum_j \sum_{\ell} \sum_i \alpha_{ji} t_{i\ell} m_j \delta(s_j) = \sum_i \sum_{\ell} t_{i\ell} p_i \mu_i^*,$$

where μ_i^* is a convex combination of $\delta(s_j)$ with $d(s_j, v_i) < \epsilon$. Thus

$$\tilde{\tilde{Q}} = \sum_{\ell=1}^{k} \left(\sum_{i=1}^{n} t_{i\ell} p_i \right) \delta(a_{\ell}),$$

where

$$a_{\ell} = \frac{\sum_j \sum_i \alpha_{ji} t_{i\ell} m_j s_j}{\sum_j \sum_i \alpha_{ji} t_{i\ell} m_j} = \frac{\sum_i t_{i\ell} p_i v_i^*}{\sum_i t_{i\ell} p_i}.$$

Note that $d(a_{\ell}, b_{\ell}) < \epsilon$. Thus $\tilde{\tilde{Q}}$ is an ϵ-perturbation of $\tilde{Q}$, and Part 1 is proved.

Part 2. V is D-dimensional (but P is not required to live on a set of finite diameter). Let $\alpha > 0$ and let $\lambda \geq 1$ be such that $\int_{\|X\| > \lambda} \|x\| \, dP(x) < \alpha/k$ and $\|y_j\| \leq \lambda$, $j = 1, \ldots, k$. Let $B = \{x\colon \|x\| \leq \lambda\}$. Let $\{e_i, \ i = 1, \ldots, 2^D\}$ be the vertices of a minimal D-cube containing B, so $\|e_i\| = \lambda\sqrt{D}$ for each i. Let $P_\alpha = P|_B + (\alpha/\lambda)\sum_{i=1}^{2^D} \delta(2e_i)$. This is close to P if α is small and lives on a set of finite diameter.

It shall be shown that P_α convexly dominates Q. For any nonnegative convex function c on V, there are affine functions a_j, $j = 1, \ldots, k$, such that $a_j(y_j) = c(y_j)$ and $a_j(x) \leq c(x)$ for all x. Let $g(x) = \max\{a_j^+, \ j = 1, \ldots, k\}$. Then g is convex and $g \leq c$ and $\int g \, dQ = \int c \, dQ$. So $\int c \, dP_\alpha \geq \int g \, dP_\alpha$, and $\int g \, dP \geq \int g \, dQ = \int c \, dQ$. Thus it is enough to show that $\int g \, dP_\alpha > \int g \, dP$, and since P and P_α agree on B, it is enough to show that $\int_{\sim B} g \, dP_\alpha \geq \int_{\sim B} g \, dP$. Thus it suffices to show that $\int_{\sim B} a^+ \, dP_\alpha \geq$

$k\int_{\sim B} a^+ \, dP$ for every affine function a such that $a(x) \geq 0$ for some x in B, because then

$$\int_{\sim B} \max\{a_j^+, j = 1, \dots, k\} \, dP_\alpha \geq (1/k) \sum_{j=1}^{k} \int_{\sim B} a_j^+ \, dP_\alpha \geq \sum_{j=1}^{k} \int_{\sim B} a_j^+ \, dP$$
$$\geq \int_{\sim B} \max\{a_j^+, j = 1, \dots, k\} \, dP$$

(note that for any finite sequence s_j of nonnegative numbers, $(1/k)\sum_{j=1}^{k} s_j \leq \max\{s_j, j = 1, \dots, k\} \leq \sum_{j=1}^{k} s_j$). Let $a(x) = \ell(x) + b$, where ℓ is linear. Choose m such that $a(e_m) = \max\{a(e_i), i = 1, \dots, 2^D\}$, and let $e = e_m$. Note that $a(e) \geq 0$, since B is a subset of the convex hull of the $\{e_i\}$, and $a(x) \geq 0$ somewhere in B by assumption. Note that $\ell(e) = \max\{\ell(e_i), i = 1, \dots, 2^D\}$ also, and $\ell(e) \geq 0$ also, since $-e$ is also a vertex of the D-cube and either $\ell(e)$ or $\ell(-e)$ would have to be ≥ 0. Now for any x not 0 in V, $\ell(x) = \ell((\lambda x/\|x\|)(\|x\|/\lambda)) \leq (\|x\|/\lambda)\ell(e)$, since $\lambda x/\|x\|$ is in B, which is a subset of the convex hull of the $\{e_i\}$. Let $A = \sim B \cap \{x\colon a(x) \geq 0\}$. Thus $\int_{\sim B} a^+(x) \, dP(x) = \int_A (\ell(x) + b) \, dP(x) \leq \ell(e) \int_A (\|x\|/\lambda) \, dP(x) + bP(A) \leq \ell(e)\alpha/(\lambda k) + bP(A)$.

Case 1. $b \geq 0$. Note $\int_{\sim B} \|x\| \, dP(x) \leq \alpha/k$, so $P(\sim B) \leq \alpha/(\lambda k)$. So $\ell(e)\alpha/(\lambda k) + bP(A) \leq (\ell(e) + b)\alpha/(\lambda k) \leq (\ell(2e) + b)\alpha/(\lambda k)$ (since $\ell(e) > 0) = a(2e)\alpha/(\lambda k) \leq (1/k)\int_{\sim B} a^+ \, dP_\alpha$.

Case 2. $b < 0$. $\ell(e)\alpha/(\lambda k) + bP(A) \leq \ell(e)\alpha/(\lambda k) \leq (\ell(e) + a(e))\alpha/(\lambda k)$ (since $a(e) \geq 0) = (\ell(e) + \ell(e) + b)\alpha/(\lambda k) = a(2e)\alpha/(\lambda k) \leq (1/k)\int_{\sim B} a^+ \, dP_\alpha$. Thus in either case

$$\int_{\sim B} a^+ \, dP \leq (1/k) \int_{\sim B} a^+ \, dP_\alpha.$$

It has now been proved that P_α convexly dominates P, hence Q.

Since P_α lives on a set of finite diameter, Part 1 of the proof yields a Borel partition A_i, $i = 1, \dots, n$, of V, and row-stochastic t_{ij}, $i = 1, \dots, n$, $j = 1, \dots, k$, such that $\mathrm{fus}(A(\ ); t(\ ,\ ); P_\alpha)$ is an α-perturbation of a finitistic measure majorizing Q. Let

$$b_j = \sum_{i=1}^{n} t_{ij} P(A_i) b(A_i, P) \bigg/ \sum_{i=1}^{n} t_{ij} P(A_i),$$

and

$$d_j = \sum_{i=1}^{n} t_{ij} P_\alpha(A_i) b(A_i, P_a) \Big/ \sum_{i=1}^{n} t_{ij} P(A_i).$$

Now

$$\left| \sum_{i=1}^{n} t_{ij} P_\alpha(A_i) - \sum_{i=1}^{n} t_{ij} P(A_i) \right| \le \sum_{i=1}^{n} \left(P_\alpha(A_i \cap \sim B) + P(A_i \cap \sim B) \right)$$

$$\le P_\alpha(\sim B) + P(\sim B) \le \alpha(2^D + 1)$$

(since λ assumed ≥ 1). Also,

$$\left\| \sum_{i=1}^{n} t_{ij} \left(P_\alpha(A_i) b(A_i, P_\alpha) - P(A_i) b(A_i, P) \right) \right\|$$

$$= \left\| \sum_{i=1}^{n} t_{ij} \left(\int_{A_i \cap \sim B} x \, dP_\alpha(x) - \int_{A_i \cap \sim B} x \, dP(x) \right) \right\|$$

$$\le \int_{\sim B} \|x\| \, dP_\alpha(x) + \int_{\sim B} \|x\| \, dP(x)$$

$$\le (\alpha/\lambda) \sum_{i=1}^{2^D} \|2 e_i\| + \alpha/k \le \alpha 2^D 2\sqrt{D} + \alpha/k.$$

Since $q_j > 0$ and D and k are fixed, and since $\sum_{i=1}^{n} t_{ij} P_\alpha(A_i) \ge q_j(1 - \alpha)$, it is clear that by taking α sufficiently small, $\operatorname{dist}(b_j, d_j) < \epsilon - \alpha$, since the numerators and denominators in the expressions for b_j and d_j can be made arbitrarily close. Thus $\operatorname{fus}(A(\ ); t(\ ,\ ); P)$ is an ϵ-perturbation of a measure majorizing Q.

Part 3. Finally, allow V to be a separable Banach space (the convex compact metrizable subset of a lctvs case is similar but even simpler, because the measure already would live on a compact set; that case is left to the reader).

Let $\alpha > 0$. Choose K, a compact subset of V, such that $P(\sim K) < \alpha$ and

$$\int_{\sim K} \|x\| \, dP(x) < \alpha,$$

and y_j is in K for $j = 1, \ldots, k$. (This can be done since P has a finite first moment). Every Banach space is isometric to a subspace of one with the 1-approximation property ([7, p. 37]), so it may be assumed that V has this property. This means there exists T a finite rank (that is, finite-dimensional range) linear operator on V such that $\operatorname{dist}(x, Tx) < \alpha$ for all x in K and $\operatorname{norm}(T) = 1$. Let P_T be the measure defined on range(T) by $P_T(A) = P(T^{-1}(A))$, and similarly for Q_T (note that $Q_T = \sum_{j=1}^{k} q_j \delta(Ty_j)$). Since T is linear, c composed with T is convex for any convex function c on the range of T, so P_T also convexly dominates Q_T. So there exists n and an $n \times k$ row-stochastic matrix (t_{ij}) of nonnegative elements and a Borel partition B_i, $i = 1, \ldots, n$, of range(T), such that $\operatorname{fus}((B_i); (t_{ij}); P_T)$ is an α-perturbation of a measure majorizing Q_T. Let $A_i = T^{-1}(B_i)$, $i = 1, \ldots, n$. Then $P(A_i) = P_T(B_i)$. Let

$$a_j = \sum_{i=1}^{n} t_{ij} P(A_i) b(A_i, P) \Big/ \sum_{i=1}^{n} t_{ij} P(A_i),$$

and

$$b_j = \sum_{i=1}^{n} t_{ij} P_T(B_i) b(B_i, P_T) \Big/ \sum_{i=1}^{n} t_{ij} P_T(B_i).$$

Now

$$\begin{aligned} \| P_T(B_i) b(B_i, P_T) - P(A_i) b(A_i, P) \| &= \left\| \int_{B_i} y \, dP_T(y) - \int_{A_i} x \, dP(x) \right\| \\ &= \left\| \int_{A_i} (Tx - x) \, dP(x) \right\| \end{aligned}$$

(using change-of-variable formula $\int y \, dP_T(y) = \int Tx \, dP(x)$). But since $\operatorname{dist}(x, Tx) < \alpha$ for x in K and $\operatorname{norm}(T) = 1$, this is $\leq \alpha P(A_i \cap K) + 2 \int_{A_i \cap \sim K} \|x\| \, dP(x)$. Thus

$$\begin{aligned} \operatorname{dist}(a_j, b_j) &\leq \sum_{i=1}^{n} \left(t_{ij} \alpha P(A_i) + 2 \int_{A_i \cap \sim K} \|x\| \, dP(x) \Big/ \sum_{i=1}^{n} t_{ij} P(A_i) \right) \\ &\leq \alpha + 2 \int_{\sim K} \|x\| \, dP(x) / q_j (1 - \alpha) \leq \alpha + 2\alpha / q_j (1 - \alpha). \end{aligned}$$

Now $\operatorname{dist}(y_j, Ty_j) < \alpha$, since y_j is in K by definition. Thus $\operatorname{dist}(y_j, a_j) \leq \operatorname{dist}(y_j, Ty_j) + \operatorname{dist}(\mathrm{T}y_j, b_j) + \operatorname{dist}(b_j, a_j) \leq \alpha + \alpha + \alpha + 2\alpha / q_j (1 - \alpha)$, which is $< \epsilon$ for sufficiently small α. ∎

In the non-Banach space case, where there is no norm, V still embeds in a space with the approximation property ([3, p. 437]), meaning that on any compact set the identity can be uniformly approximated by a continuous linear operator of finite rank, without any global statement about the behavior of T off the compact set (analogous to the condition norm$(T) = 1$ that was used in the Banach space case). Since in this case P already lives on a compact set, no such condition is needed, and the proof is even easier.

LEMMA 4.8. *Let V, P, and Q be as in Lemma* 4.7. *Then there is a fusion of P that majorizes Q.*

Proof. By Lemma 4.7, for each n there is fusion P_n of P that is a $(1/n)$-perturbation of a finitistic measure majorizing Q. The set of fusions of P is tight ([3, p. 435]), so some subsequence of (P_n) converges weakly, to a measure that obviously majorizes Q, and is a fusion of P, since the set of fusions of P is weakly closed. ∎

Proof of Theorem 4.2. It is elementary that Q is the weak limit of a sequence Q_n of finitistic measures such that Q (hence P) convexly dominates Q_n (just take a partition into subsets of diameter $< 1/n$ for a compact set on which all but $(1/n)$ of the mass of Q lives, and collapse each set in the partition to its barycenter). Each Q_n is majorized by a fusion of P_n, by Lemma 4.8. Some subsequence of (P_n) converges to a fusion P^* of P since the set of fusions of P is tight, and this measure obviously majorizes Q (proof: for any bounded, continuous, nonnegative function f, $\int f\,dP^* = \lim \int f\,dP_n \geq \lim \int f\,dQ_n = \lim \int f\,dQ$; note that P^* majorizes Q iff $\int f\,dP^* \geq \int f\,dQ$ for all bounded continuous nonnegative functions f, and recall that P_n converges weakly to P^* iff $\lim \int f\,dP_n = \int f\,dP^*$ for all bounded continuous f). ∎

ACKNOWLEDGMENT

The authors are indebted to Prof. Y. Rinott for pointing them to reference [4] below.

REFERENCES

1. D. Blackwell, Equivalent comparisons of experiments, *Ann. Math. Statist.* **24** (1953), 265–272.
2. P. Cartier, J. Fell, and P. Meyer, Comparison des mesures portées par un ensemble convexe compact, *Bull. Soc. Math. France* **92** (1964), 435–455.
3. J. Elton and T. Hill, Fusions of a probability distribution, *Ann. Probab.* **20** (1992), 421–454.
4. P. Fischer and J. Holbrook, Balayage defined by the nonnegative convex functions, *Proc. Am. Math. Soc.* **79** (1980), 445–448.
5. G. Hardy, J. Littlewood, and G. Pólya, Some simple inequalities satisfied by convex functions, *Messenger Math.* **58** (1929), 145–152.

6. G. Hardy, J. Littlewood, and G. Pólya, Inequalities, Cambridge Univ. Press, Cambridge, 1959.
7. J. Lindenstrauss and L. Tzafriri, "Classical Banach Spaces I," Springer-Verlag, Berlin/New York, 1977.
8. L. Mirsky, Majorization of vectors and inequalities for convex functions, *Monatsh. Math.* **65** (1961), 159–169.
9. R. Phelps, "Lectures on Choquet's Theorem," Van Nostrand-Reinhold, New York, 1966.
10. V. Strassen, The existence of probability measures with given marginals, *Ann. Math. Statist.* **36** (1965), 423–439.